SesameStreet++

James Rowell

February 2, 2017

There are 10 sorts of people in the world: those who understand binary and those who don't.

Most of us think about "whole numbers" not too differently from the way we learned to count by watching Sesame Street, the difference being that now we can count a little higher. How we've trained ourselves, it's automatic to think the way that we write a number or say a number *is* the number.

If I owe you 13 cents and I give you one dime and three pennies then after thanking me profusely for repaying this staggering debt, we'll agree that it's settled with those coins equaling 13 pennies. We identify the symbol "13" very strongly with this particular number - it would be tough to get through life in the modern world without such an automatic process running in our brains. This example highlights what this particular symbol "13" actually means - one dime (1×10) plus three pennies (3×1).

Let's look at the number 13 in some alternative ways - it's the number of months in a year plus one month; what I'm suggesting is that there is no need for the symbol "13" in order to think about this particular number of months. Similarly, 13 is this many apples 🍎🍎🍎🍎🍎🍎🍎🍎🍎🍎🍎🍎🍎; or 13 is the sixth prime number. None of these ways of thinking about the number 13 require that we represent it using the digits 1 and 3 butted up next to each other.

Each number exists independently from any symbol or word that might represent it. Numbers are an idea - perhaps such a strong idea that the universe wouldn't exist without them! Anyway, for our purposes whole numbers exist in some abstract realm - Each number is one whole unit more than the previous number, starting at nothing, that is "zero", and jumping to something, that is "one", then one more, which gets us to "two", then again to "three", etc. Continuing in this way forever... we get them all.

To get back to the idea of what a whole number really is, try to forget about the symbols or words we use and picture a pile of apples. There are zero apples (it's hard to show no-apples), then we introduce an "🍎" to get our very first, and smallest, non-empty pile of apples. Then add another apple to get a pile of "🍎🍎", then "🍎🍎🍎", then "🍎🍎🍎🍎" then some big pile of "🍎🍎🍎🍎🍎...🍎🍎🍎" after we've been adding apples for a while. Each successively bigger pile of apples corresponds with each successive whole number.

We expand this entire set of whole numbers to include their negative-counterparts and call this larger set "integers". We denote the set of integers with this symbol: $\mathbb{Z}$.

However, using a "1" followed by a "3" to represent the integer "🍎🍎🍎🍎🍎🍎🍎🍎🍎🍎🍎🍎🍎" is VERY handy. So we use Hindu-Arabic numbers and the positional notation of "base-ten", more commonly known as "decimal", to represent each specific integer. We slap a "-" on the front if we need to talk about a negative integer.

Base-ten representation of an integer is far superior to ancient Roman numerals for example. Try adding two numbers together in ancient Rome, or worse, multiplying or dividing them. What's XI times IX? Would you believe me if I told you it's XCIX? Unless you convert those to Hindu-Arabic numerals to check, you're just gonna have to trust me. Truth is - I don't know how to multiply using Roman numerals - nor did most Romans! Not only that, but I'll bet that most kids who graduated from Sesame Street can count higher than any Roman could - as the Roman system only effectively allowed counting up to 4999.

Using base-ten for us is automatic, we barely think about it when we're adding numbers or multiplying them - but it's worth looking carefully at how base-ten works - so let's examine it from the ground up*.

It's useful to have simple symbols to represent each of the integers from one to nine, namely our familiar 1, 2, 3, 4, 5, 6, 7, 8 and 9 which have an interesting history and predate their use in base-ten.

Slightly more modern, but still quite ancient, is the symbol "0" for "zero", originally meaning "empty". Zero also predates its use in base-ten but without zero, base-ten wouldn't be possible.

Base-ten uses the idea of stringing a series of digits together (a digit being one of the numbers 0, 1, 2, ... 9), one after the other to be able to represent each whole number. Let's look at the first two-digit number, that is, ten, which as you well know looks like this: "10". This extra digit on the left tells us how many tens we have and the last, or rightmost digit says how many additional units to add to it.

So our very first two-digit number 10 means "one lot of ten - plus zero units". When we see "11" - we interpret it to mean "one lot of ten - plus one unit", and "12" is "one lot of ten - plus two units", etc. Continuing on; "20" - we interpret to mean "two lots of ten, plus zero units", etc. up to "90" meaning "nine lots of ten, plus zero units".

Following this line of reasoning since "10" now means the integer ten, then "100" must mean "ten lots of ten, plus zero units"- which is exactly what it means. We have a special word for this number we call it "one hundred" or "one lot of a hundred, plus zero lots of tens, plus zero units". Similarly "200" means "two lots of a hundred, plus zero lots of ten, plus zero units", etc.

We can keep going by one-hundred until we similarly get to "1000" or "ten lots of a hundred, plus zero lots of ten, plus zero units" otherwise known as "a thousand" or more specifically "one lot of a thousand, plus zero lots of a hundred, plus zero lots of ten, plus zero units".

It gets a little tedious to be so specific when reading out a number so our language has developed quite a few verbal shortcuts. Furthermore it doesn't take long before we run out of fancy names for these "powers-of-ten" like, million, billion, trillion, zillion etc. So let's introduce some nice clean mathematical notation to describe these powers-of-ten and let's forget the fancy words.

*Please forgive the incredibly obvious nature of much of the following discussion, but I want to take a good running start at some more unfamiliar notions. Perhaps looking at the familiar with fresh eyes will help in seeing the new ideas easier.

$$\begin{aligned}
100 &= 10 \times 10 = 10^2, \\
1000 &= 10 \times 10 \times 10 = 10^3, \\
10000 &= 10 \times 10 \times 10 \times 10 = 10^4, \\
&\dots \\
\underbrace{10\dots000}_{\text{k zeros}} &= \underbrace{10 \times 10 \times 10 \times 10 \times \dots \times 10}_{\text{k 10s}} = 10^k
\end{aligned}$$

10^k means there are k tens multiplied together - also written as a 1 followed by k zeros[*]. The above list explicitly shows the cases for $k = 2, 3$ and 4. Using the k like that is just a way to show that we can pick ANY whole number, i.e., there is no limit on how big k can be.

The notation of 10^k is very handy, in fact it extends to the case when $k = 0$ and $k = 1$.

So 10^1 means[†] that there is only one ten multiplied together, or one "0" following the "1", in other words just the number ten itself.

How about when $k = 0$? Examining the pattern of how the power k relates to how many zeros follow the "1" (eg, $10^1 = 10$, $10^2 = 100$, $10^3 = 1000$, etc.) then it must be the case that $10^0 = 1$, i.e., no zeros follow the "1", which is exactly right. Furthermore every number raised to the 0^{th} power is 1.[‡]

Let's look at an example. Reading the number 92507 out according to our technique we can see that it's "nine lots of ten-thousand, plus two lots of a thousand, plus five lots of a hundred, plus zero lots of ten, plus seven units":

$$\begin{array}{rrcr c rr}
 & 9 & \times & 10000 & & & 90000 \\
+ & 2 & \times & 1000 & & + & 2000 \\
+ & 5 & \times & 100 & = & + & 500 \\
+ & 0 & \times & 10 & & + & 00 \\
+ & 7 & \times & 1 & & + & 7 \\
\cline{6-7}
 & & & & & = & 92507
\end{array}$$

Written[§] in terms of powers-of-ten: $92507 = 9{\times}10^4 + 2{\times}10^3 + 5{\times}10^2 + 0{\times}10^1 + 7{\times}10^0$.

You can think of each digit as being a little dial that controls how many lots of its corresponding power-of-ten will contribute to the value of the integer. Please note that we went ahead and multiplied the units digit (7 in our example) by 10^0 (which is the same as multiplying by 1) so that we can see that there's nothing special about the units digit, it's just some number between 0 and 9, times a power-of-ten like each of the other digits. It also adds a kind of beauty to the expression in that all the digits can be treated in a similar matter.

[*] k is called the "exponent" and you should read the symbol 10^k as "ten-raised-to-the-k^{th}-power" or "ten-to-the-k", so 10^2 is "ten raised to the second power" or 10^4 is "ten-to-the-fourth". You may also see 10^2 referred to as "ten squared", similarly 10^3 as "ten cubed" - but since we don't live in 4 dimensional hyperspace, we don't have a way of saying 10^4 that has geometric meaning.

[†] Don't forget to read 10^1 as "ten-to-the-one".

[‡] Proof: Since $a^{b+c} = a^b a^c$ consider when $c = 0$; that is, $a^b = a^{b+0} = a^b a^0$ so because of the uniqueness of the multiplicative identity "1", then a^0 *must* be 1 since it's behaving like a "1" in the expression $a^b = a^b a^0$.

[§] Recall the mnemonic "bedmas" for the "Order of Operations" in evaluating an expression, which is no different from what we did in our table above the expression.

Claim: Given that we can use as high a power-of-ten as we like and we can string together *as long a list of digits* as pleases us, that means that we can create *any* positive integer we want no matter how big it is.

That's a pretty tall claim.

How do we know that we can create ALL the positive integers with this scheme? For example, how do we know that we didn't miss one? Or how do we know that some string of digits doesn't represent two different integers? I know it seems silly to ask these kinds of questions - after all, people have been counting in base-ten for almost two thousand years, if there was a problem, you'd think we'd have heard about it by now! ...so, obviously it works.

Here's the thing about mathematics - the *only* ideas we take as obvious are the axioms - those are the mathematical ideas that are so simple that they can't be expressed in yet other even-simpler ideas. The axioms are the minimal set of simple, obvious, irrefutable ideas from which everything else in mathematics is built[*].

As obvious as it seems, the fact that we can use base-ten to represent the integers is NOT among the list of axioms.

As discussed in the opening paragraph, we think that the way that we write a number, or say a number *is* the number - no problem - it's totally fine to think this way. Numbers written in base-ten are in a perfect one-to-one correspondence with the integers so it's safe to think about numbers written in base-ten as the integers.

However, the reason that we *really* know that it's safe to think this way is because we spelled out what it means to represent a number in base-ten with a theorem, then we've proven that the theorem is true. A proof is just a series of arguments that logically connects our theorem directly[†] to the axioms, so that the only way that our theorem could be false is if the axioms themselves are false.

So let's think about what a "Base-Ten Representation Theorem" might look like so that once we prove it we may confidently go forward using base-ten with impunity knowing that we can represent each and every integer with it's own unique label.

For starters, we intuitively know that counting with base-ten covers all the possible integers. After all, how we count leaves no room for gaps. We know that if we count up to say... 499, that if we add 1, the 9's roll over to 0, and 1 gets added to the 4, like an odometer in your car, getting us to the very next integer which is 500. We know this always works no matter what number we add 1 to.

Our intuition is good - let's write it down as something that has to be in our theorem. We might say something like:

Base-Ten Representation Theorem (initial draft)

Every integer has a representation in base-ten.

[*] The axioms: For every integer a, b and c: Associativity: $(a+b)+c = a+(b+c)$ and $a(bc) = (ab)c$; Commutativity: $a+b = b+a$ and $ab = ba$; Distributive: $a(b+c) = (b+c)a = ab+ac$; Identities: There are integers 0 and 1 such that, $a+0 = 0+a = a$ and $a \cdot 1 = 1 \cdot a = a$ and Additive Inverse: $a+(-a) = 0$. Note: in general integers do NOT have multiplicative inverses that are also integers. (eg. $\frac{1}{2}$ is the multiplicative inverse of 2 because $\frac{1}{2} \cdot 2 = 1$ but $\frac{1}{2}$ is not an integer!)

[†] directly ...or indirectly via other previously proven theorems.

Something else we know intuitively is that each number written in base-ten represents only ONE integer. It's almost feels silly to spell it out, but if we were to count out 4 piles of 100-apples-each-pile, then 9 piles of 10-apples-each-pile, then count out 9 additional apples, *then* scoop them all into a big pile that we'd always get the exact same size big-pile-of-apples.

It goes the other way too. If we were handed the aforementioned big pile of apples we could start counting out piles of 100. We'd try to make as many piles of 100 as we could, and we'd find that we'd have 4 piles of 100 before we couldn't make another such pile. Then we would start counting out piles of 10 with the remaining apples. After we made as many piles of ten as we could out of those remaining apples, we would discover that we'd have 9 such piles of ten with 9 apples left over, in other words 499 apples! There is NO other way to divvy up this big pile of apples if we follow this procedure. In other words, each integer is represented by only ONE base-ten number.

Let's strengthen our theorem based on the last two observations.

Base-Ten Representation Theorem (second draft)

Every integer has a unique representation in base-ten.

We know that for every non-zero integer, there exists another unique integer such that if you add them together you get zero, or algebraically:

$$\text{For all integers } a \neq 0 \text{ there exists a unique integer } -a \text{ such that } a + (-a) = 0$$

This means that EVERY positive integer has a unique negative counterpart. So if we can prove the Base-Ten Representation Theorem for the positive integers then we can easily extend it to the negative integers with a corollary. Our corollary will say something like "Slap a '-' (minus) sign in front of the positive base-ten representation to get a unique representation for it's negative counterpart" ...or some such words. Also zero is easy to deal with separately, so to keep our lives simple let's restrict our attention to only proving the theorem for positive integers.

Moving forward it would be helpful to have a name for our positive integer so that we can refer to it directly - how about n for "number":

Base-Ten Representation Theorem (third draft)

Every positive integer n has a unique representation in base-ten.

At the moment it's not very helpful to have named n (the theorem as it stands doesn't say anything more about n so why did we bother naming it?) but as we flush out the remaining details of the theorem we can refer to n which carries the important information that it could be ANY positive integer.

Back a few pages we looked at the real meaning of the number 92507 by adding together various lots of powers-of-ten[*]:

$$92507 = 9{\cdot}10^4 + 2{\cdot}10^3 + 5{\cdot}10^2 + 0{\cdot}10^1 + 7{\cdot}10^0$$

Every base-ten number implicitly describes an algebraic expression like this, so let's come up with a general expression of this form that can describe ANY positive integer n.

[*] It's time to replace our "×" symbol for multiplication, with "·" because "×" might get confused for an "x" in an expression, whereas "·" never will be. Eg. $x \times 2$ vs. $x \cdot 2$, additionally ending up with something that is more aesthetically pleasing. You may also see the "·" omitted entirely as in ab - which means $a \cdot b$ as you have seen in earlier footnotes.

It's important to remember that we only allow n's digits to take on the values 0 through 9 - no other values are allowed. Most noteworthy is there are at most 9 lots of each given power-of-ten in our base-ten-representation of n. For example, if we counted out 9 piles of ten-apples each, but still ten more apples to count, we would simply add that extra pile of ten-apples to the other 9 piles of ten-apples and count it as 1 pile of a hundred-apples. That's the rule - we are only allowed at most 9 lots of a certain power-of-ten sized pile-of-apples.

Let the term d represent one of the digits of n. So d represents how many lots of some power-of-ten contribute to the value of n. d must be an integer between 0 and 9 inclusive (we could say "0 is less than or equal to d which is also less than or equal to 9" or better "$0 \leq d \leq 9$"), so the contribution of the digit d to the value of n would be this amount:

$$d \cdot 10^i, \text{ where } d \text{ is an integer such that } 0 \leq d \leq 9 \text{ and } i \geq 0$$

In other words, for some power-of-ten (that's the i which can be anything, like 0, 1, 2, 3, etc.) we have d lots of it. Note that d could be 0, which means that it doesn't contribute ANY lots of 10^i to the value of n, all the way up to the maximum contribution of $9 \cdot 10^i$.

Let's take a moment to consider why it's important to let people know that d is an integer. From the context of our discussion you and I know we're only talking about 0, 1, 2, 3, 4, 5, 6, 7, 8 or 9 when we say that d is in the range $0 \leq d \leq 9$. However, if someone were to come upon the following statement in the wild:

$$d \cdot 10^i, \text{ such that } 0 \leq d \leq 9$$

...how would they know that d isn't a fraction like $\frac{22}{7}$, or the value of π or something weird like that? Answer: They wouldn't know - which is why it's important to spell it out.

Our general expression for n should only have one integer digit in the range of 0 to 9 associated with each power-of-ten, in other words our expression should never allow this:

$$4 \cdot 10^2 + 7 \cdot 10^2$$

because we could group those powers-of-ten together like this:

$$4 \cdot 10^2 + 7 \cdot 10^2 = (4+7)10^2 = 11 \cdot 10^2$$

...meaning the value of the *actual* digit associated with 10^2 is eleven which violates our rule of keeping the value of the digit in the range of 0 to 9.

n is likely composed of *many* digits, each of which is multiplied by a distinct power-of-ten. Suppose n is a five digit number. We've been using the specific five-digit number 92507 in our examples above, but now suppose n is some number from 10000 all the way up to 99999 (which is a complete list of all the five-digit numbers). If we were to describe this general five digit number, then we would need five different 'd's for each of the five digits. Somehow we need to associate a different term d with each of the powers 10^4, 10^3, 10^2, 10^1 and 10^0.

Mathematics has a convention for coming up with a list of terms for situations just like this - we tack a subscript onto the name like so: d_2 which you read as "dee-two"[*]. d_2 is a term to represent a digit just like the d we used above. But now we can use that little subscript as a way to associate it to a specific power-of-ten. Naturally we'll associate d_2 with 10^2 (ten squared) as follows; d_2 contributes to the value of n by this amount:

$$d_2 \cdot 10^2, \text{ where } d_2 \text{ is an integer such that } 0 \leq d_2 \leq 9$$

[*] ...yes like artoo-detoo, which perhaps George should have writen as "R_2D_2" and not "R2-D2"!

If we define d_2 like this, then we know that when we refer to the digit d_2 that we are talking about the digit that is multiplied with 10^2.

Let our five-digit-number n use d_0, d_1, d_2, d_3 and d_4 for its digits. Then the general expression for n looks like this:

$$n = d_4{\cdot}10^4 + d_3{\cdot}10^3 + d_2{\cdot}10^2 + d_1{\cdot}10^1 + d_0{\cdot}10^0$$

Since we're getting something that looks like hard-core math, let's take a moment to read it out loud, as a Math-Professor might do in a lecture. It might sound like this:

> "n is equal to dee-four times ten-to-the-fourth, ... plus dee-three times ten-cubed, ... plus dee-two times ten-squared, ... plus dee-one times ten, ... plus dee-zero times one."

When we write out our expression for our five-digit number n we want to be sure to let people know that each of the digits d_0, d_1, d_2, d_3 and d_4 are integers and must be between 0 and 9 inclusive.

Furthermore, to make sure n is a legitimate five-digit number we have to call out the exception that d_4 can NOT be zero - it has to be at least 1. Why? Because if d_4 were zero then n would only be a four-digit number, or perhaps a three-digit number, or only two-digits etc.

Let's call out this important information, along with the expression for n:

$$\begin{aligned} n = d_4{\cdot}10^4 + d_3{\cdot}10^3 + d_2{\cdot}10^2 + d_1{\cdot}10^1 + d_0{\cdot}10^0&\text{, where} \\ &d_0 \text{ is an integer such that } 0 \le d_0 \le 9\text{, and} \\ &d_1 \text{ is an integer such that } 0 \le d_1 \le 9\text{, and} \\ &d_2 \text{ is an integer such that } 0 \le d_2 \le 9\text{, and} \\ &d_3 \text{ is an integer such that } 0 \le d_3 \le 9\text{, and} \\ &d_4 \text{ is an integer such that } 1 \le d_4 \le 9. \end{aligned}$$

Ok, hold on a minute - that's getting a little cumbersome; plus did you notice how we slipped in that different range for d_4? Not a great presentation.

A common convention in situations like this is to let another term like i, for perhaps "index", stand in for the subscript when you want to talk about all your 'd's at once:

> Let d_0, d_1, d_2, d_3 and d_4 be integers such that:
>
> $$n = d_4{\cdot}10^4 + d_3{\cdot}10^3 + d_2{\cdot}10^2 + d_1{\cdot}10^1 + d_0{\cdot}10^0$$
>
> where $0 \le d_i \le 9$ for all i in $\{0, 1, 2, 3, 4\}$ and $d_4 \ne 0$.

Better! That expression for n describes every five-digit number - now we need to extend our hard-coded five-digit expression for n to an arbitrary number of digits. Consider the following progression:

expression for n	number-of-digits
$d_0{\cdot}10^0$	1
$d_1{\cdot}10^1 + d_0{\cdot}10^0$	2
$d_2{\cdot}10^2 + d_1{\cdot}10^1 + d_0{\cdot}10^0$	3
$d_3{\cdot}10^3 + d_2{\cdot}10^2 + d_1{\cdot}10^1 + d_0{\cdot}10^0$	4
$d_4{\cdot}10^4 + d_3{\cdot}10^3 + d_2{\cdot}10^2 + d_1{\cdot}10^1 + d_0{\cdot}10^0$	5
...	...
$d_k{\cdot}10^k + \ldots + d_4{\cdot}10^4 + d_3{\cdot}10^3 + d_2{\cdot}10^2 + d_1{\cdot}10^1 + d_0{\cdot}10^0$	k+1

Using k like this let's us specify any number of digits we want. If we let $k = 0$ we get the first "single digit" item on the list. $k = 4$ gives us our five-digit number above, or we could let k be a billion, which would allow us to specifcy an integer that has a billion-and-one digits in it[*]!

So there we have it, we found our expression for being able to express each positive integer, let's use it in a revised draft of our theorem:

Base-Ten Representation Theorem (close to final draft)

For every positive integer n there is a unique sequence of integers $d_0, d_1, d_2, \ldots, d_k$ such that:

$$n = d_k{\cdot}10^k + d_{k-1}{\cdot}10^{k-1} + \cdots + d_2{\cdot}10^2 + d_1{\cdot}10^1 + d_0{\cdot}10^0$$

where $0 \leq d_i \leq 9$ for all i in $\{0, 1, 2, \ldots, k\}$ and $d_k \neq 0$.

Definition: n is represented in base-ten by the string of digits $d_k d_{k-1} {\cdots} d_2 d_1 d_0$

Our newly added "Definition" introduces exactly what it means to write the number out in base-ten; that is, we toss out all the extraneous stuff from our expression and string all the digits one after another. Starting at the most-significant digit d_k on the left, down to the next digit to its right which is d_{k-1} (read as "dee-kay-minus-one"[†]) all the way down to the least-significant units-digit d_0 on the right.

We are so darn close, but there is one picky detail that we have to be careful about. We can safely use the symbols for the integers from *zero* to *nine*, as they exist independently from base-ten[‡]. *However* this theorem defines what base-ten means so how can we use "10" in our theorem?! The answer is, we can't - at least not without acknowledging it somehow.

We have a couple of choices. Since we don't have an actual symbol for the integer ten, we have to make one up. The first possibility is:

Let T represent the integer ten.

For every positive integer n there is a unique sequence of integers $d_0, d_1, d_2, \ldots, d_k$ such that:

$$n = d_k{\cdot}T^k + d_{k-1}{\cdot}T^{k-1} + \cdots + d_2{\cdot}T^2 + d_1{\cdot}T^1 + d_0{\cdot}T^0$$

... etc.

The other choice is we could cheat a little, and define the symbol for *ten* to be "10". I favor this approach for now, don't worry, it's not breaking any rules, you can go ahead and define things however you like in mathematics so long as it doesn't create inconsistencies. Anyway, we're not going to prove the Base-Ten Representation Theorem - we're going to keep digging, and eventually prove something more general where this problem disappears. But let's write it out anyway, since we've worked so hard to get this far PLUS it looks so nice!

[*] A billion-and-one digit number is *ridiculously* large, consider that our estimate of the number of molecules in the entire universe would only need a base-ten number with the k set to somewhere between 78 and 82 to count them all.

[†] ... and d_{k-1} is multiplied by "ten-to-the-power-of-(kay-minus-one)".

[‡] as we discussed earlier these symbols: 0, 1, 2, 3, 4, 5, 6, 7, 8 and 9 predate base-ten - they stand for the integers that represent our successive sized piles of apples from none, to 🍎 up to 🍎🍎🍎🍎🍎🍎🍎🍎🍎.

Base-Ten Representation Theorem

Let 10 represent the integer ten.

For every positive integer n there is a unique sequence of integers $d_0, d_1, d_2, \ldots, d_k$ such that:

$$n = d_k{\cdot}10^k + d_{k-1}{\cdot}10^{k-1} + \cdots + d_2{\cdot}10^2 + d_1{\cdot}10^1 + d_0{\cdot}10^0$$

where $0 \le d_i \le 9$ for all i in $\{0, 1, 2, \ldots, k\}$ and $d_k \ne 0$.

Definition: n is represented in base-ten by the string of digits $d_k d_{k-1} \cdots d_2 d_1 d_0$

Consider that base-ten is not the only base in use these days. Since the introduction of the EDVAC* computer, around 1950, there have been many orders of magnitude more calculations done in base-two (otherwise known as binary) by computers than have EVER been done by people in base-ten for the entirety of human history. (This might even be true if we only count one-day's worth of binary computer calculations - someone needs to check this.)

Binary-computer logic gates (the building blocks of the modern computer) can only take one of two states, that is; "off" or "on". We interpret these two states to represent these two numbers: 0 and 1. By doing so, in the same way that base-ten uses ten numbers 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9 for its digits; we can represent integers in base-two with just the digits 0 and 1. How is this possible? Let's find out with an imaginary trip into space.

Consider distant Planet-Nova on which the emergent intelligent species only have nine fingers on their hands. They have three hands with three fingers each - anyway, that's why they use base-nine, so they only need the numbers 0, 1, 2, 3, 4, 5, 6, 7 and 8 for their digits[†]. So like we Earthlings do for the integer ten, instead of making up a new symbol for nine, they use "10" to represent the integer nine - which for them means "One lot of nine, plus zero units".

Similarly on Planet-Ocho, since they only have eight fingers, then they use base-eight and only use numbers 0, 1, 2, 3, 4, 5, 6 and 7 for their digits. For them "10" means "One lot of eight, plus zero units".

On and on past Planet-Gary-Seven, and Planet-Secks, Planet-Penta, . . .

Finally we come upon Planet-Claire (well someone has to come from Planet-Claire, I know she came from there), where the poor blighters only have two fingers so they only use the digits 0 and 1 and base-two, so for them "10" means "one lot of two and zero units". So on Planet-Claire "10" means two. Recall above how we arrived at our 100 in base-ten, being "ten lots of ten, plus zero units" - similarly on Planet-Claire "100" in base-two for them means "Two lots of two plus zero units" in other words, four! What is "11" in base-two? Using our technique to describe the digits we see that it's "One lot of two, plus one unit", in other words three.

[*] You might be thinking, don't you mean ENIAC which was earlier? Actually no - the ENIAC used base-ten accumulators, not binary!

[†] Digit is another word for finger! Of course that's where the math term got its start.

Here's how they count on Planet-Claire using base-two:

base-two	base-ten
0	0
1	1
10	2
11	3
100	4
101	5
110	6
111	7
1000	8
1001	9
1010	10
1011	11
1100	12 (cont...)
(...cont)	
1101	13
1110	14
1111	15
10000	16
10001	17
...	
11111	31
100000	32
...	
1000000	64
10000000	128
100000000	256

Note something interesting in the list above - the powers of two, written in base-two, resemble our powers of 10 in base-ten! That is:

$$\begin{aligned} 1 &= 2^0 = 1, & 32 &= 2^5 = 100000_{\text{(base-2)}}, \\ 2 &= 2^1 = 10_{\text{(base-2)}}, & 64 &= 2^6 = 1000000_{\text{(base-2)}}, \\ 4 &= 2^2 = 100_{\text{(base-2)}}, & 128 &= 2^7 = 10000000_{\text{(base-2)}}, \\ 8 &= 2^3 = 1000_{\text{(base-2)}}, & 256 &= 2^8 = 100000000_{\text{(base-2)}}, \\ 16 &= 2^4 = 10000_{\text{(base-2)}}, & & \ldots \end{aligned}$$

Let's look at the binary number 11010 for example. Using our wordy technique to describe the number we can see that it's "One lot of sixteen, plus one lot of eight, plus zero lots of four, plus one lot of two, plus zero units":

$$\begin{array}{rrcr} & 1 & \times & 10000 \\ + & 1 & \times & 1000 \\ + & 0 & \times & 100 \\ + & 1 & \times & 10 \\ + & 0 & \times & 1 \end{array} \quad = \quad \begin{array}{rrl} & 10000 & (16) \\ + & 1000 & (8) \\ + & 000 & \\ + & 10 & (2) \\ + & 0 & \\ \hline = & 11010 & (26) \end{array}$$

Written in terms of powers of two: $11010_{\text{(base-2)}} = 26 = 1 \cdot 2^4 + 1 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2^1 + 0 \cdot 2^0$.

Each digit in base-two can be thought of as a little switch that turns on or off the contribution of its corresponding power of two.

Claim: Given that the inhabitants of Planet-Claire can use as high a power of two as they like, and that they can string together as LONG A LIST of binary-digits as pleases them, that means that they can create ANY INTEGER THEY WANT no matter how big it is.

Sound familiar? Let's restate our theorem for base-ten but rewritten for base-two.

Base-Two Representation Theorem

For every positive integer n there is a unique sequence of integers $d_0, d_1, d_2, \ldots, d_k$ such that:

$$n = d_k 2^k + d_{k-1} 2^{k-1} + \cdots + d_2 2^2 + d_1 2^1 + d_0 2^0,$$

where $0 \leq d_i \leq 1$ for all i in $\{0, 1, 2, \ldots, k\}$ and $d_k \neq 0$.

Definition: n is represented in base-two by the string of binary-digits $(d_k d_{k-1} \cdots d_2 d_1 d_0)_2$

Our new Base-Two Representation Theorem introduced some helpful new notation. How do you know what I'm talking about if I just write "1000"? Do I mean 10^3 or 2^3? If there is any possibility for confusion we write the number like this $(1000)_{10}$ for the base-ten version meaning one-thousand and $(1000)_2$ for the binary version meaning eight. That's what the "Definition" is spelling out with the "$(\ldots)_2$" extra notation.

As is hinted by the habits of our various alien friends above it seems that we can use ANY integer greater than or equal to 2 as a base (base-one doesn't really make sense - think about it for a while). In fact computer graphics artists are known to stumble upon numbers written in hexadecimal (usually relating to specifying a color-channel), which is base-sixteen.

Base-sixteen introduces some new single-character symbols to the usual numbers 0, 1, 2, thru 9, to represent the numbers 10, 11, 12, 13, 14 and 15. Base-sixteen adds the digits A, B, C, D, E and F where $\text{A}_{16}=(10)_{10}$, $\text{B}_{16}=(11)_{10}$, $\text{C}_{16}=(12)_{10}$, $\text{D}_{16}=(13)_{10}$, $\text{E}_{16}=(14)_{10}$, $\text{F}_{16}=(15)_{10}$. So $(\text{80FB})_{16}$ is a four digit number in base-sixteen. (As we'll see shortly it means $(33019)_{10}$ in base-ten).

Note that if we omit the parentheses and subscript from a number, it means we're talking about it in base-ten; our "default" base. Case in point: the subscripts that we use to denote the base (like the "16" in $(\text{80FB})_{16}$) are written in base-ten!

We could go ahead and prove our "Base-Ten" and "Base-Two" theorems above, but what about proving the "Base-Nine" version of the theorem for the aliens on Planet-Nova, or the "Base-Eight" version for the inhabitants of Planet-Ocho?

To cover all bases (pun intended) let's restate our theorem for the general case, call it "base-b", where b is some number greater than or equal to two. If we can prove that theorem, then we'll automatically get all the cases of specific bases for free.

Basis Representation Theorem

Let b be a positive integer greater than 1.

For every positive integer n there is a unique sequence of integers $d_0, d_1, d_2, \ldots, d_k$ such that:

$$n = d_k b^k + d_{k-1} b^{k-1} + \cdots + d_2 b^2 + d_1 b^1 + d_0 b^0,$$

where $0 \leq d_i \leq (b-1)$ for all i in $\{0, 1, 2, \ldots, k\}$ and $d_k \neq 0$.

Definition: n is represented in base-b by the string of base-b-digits $(d_k d_{k-1} \cdots d_2 d_1 d_0)_b$

So to get the "Base-Ten Representation Theorem" let b equal ten. To get the "Base-Two Representation Theorem" let $b = 2$; or the "Base-Nine Representation Theorem" let $b = 9$; etc.

Bonus: Because of the general nature of the "Basis Representation Theorem" we also know that we can safely convert between different bases. (Why? ...exercise left for the student). Recall how we defined $(A)_{16} = 10$ and $(F)_{16} = 15$ as base-sixteen digits, then:

$$9,937,906 = 9 \cdot 16^5 + 7 \cdot 16^4 + 10 \cdot 16^3 + 3 \cdot 16^2 + 15 \cdot 16^1 + 2 \cdot 16^0 = (97A3F2)_{16}$$

Epilogue

At this point we really ought to get to the proof of the "Basis Representation Theorem".

However, truth be told, the proof supplied below is a little heavy going - not difficult, but not really a great introductory proof for the first time mathematician. In fact here's what one of the world's greatest mathematicians[*], G. H. Hardy, had to say about such proofs[†]:

> "We do not want many 'variations' in the proof of a mathematical theorem: 'enumeration of cases', indeed, is one of the duller forms of mathematical argument. A mathematical proof should resemble a simple and clear-cut constellation, not a scattered cluster in the Milky Way."

I couldn't agree more which is why I'm pushing the proof to a section at the end of this paper, which you may feel free to skip. I still invite you to to take a stab at following it, I tried my best to make it clear and interesting, but don't feel bad if it makes your head hurt. If you do try to follow it - it does have some points of interest - not the least of which is that it works - perhaps not as elegantly as Mr. Hardy could provide to us, but I'm no G. H. Hardy.

I think at this point you are well equipped you to try your hands at a couple of exercises for fun. If you get stuck, or to check your work, the answers are also supplied below - but please don't peek until you try the questions yourself!

Exercises

1. What are the following numbers expressed in base-ten?

   i) $(110101)_2$

   ii) $(A053D)_{16}$

   iii) $(1017)_{23}$

2. What are the following base-ten numbers expressed in an alternate base?

   i) 33 expressed in base-two?

   ii) 127 expressed in base-two? (Hint: $127 = (128 - 1)$)

   iii) 8079 expressed in base-sixteen?

   Hint: For a moment, pretend that we don't use base-ten to write out our numbers, instead picture a pile of apples. Can you picture 7654 apples? Yes? Good let's use 7654 as our example.

[*] along with the Gauss and Euclid mentioned later in this paper.

[†] in an essay he wrote called "A Mathematician's Apology".

Let's divide 7654 by 10 so we get the following:

$$7654 = 765 \cdot 10 + 4$$

Notice the remainder 4 is the least significant digit of our integer 7654 (i.e. the d_0 digit in the theorem).

How do we get the next digit, i.e. the d_1 digit that corresponds to the 10^1 term? Well, it's kind of cheating, but since we happen to be looking at that last expression written in base-ten we can see it sitting right there in at the end of the quotient "765". So, let's use the same technique and divide 765 by 10:

$$765 = 76 \cdot 10 + 5$$

So the remainder is 5 our d_1 digit. Let's keep going, this time dividing the previous quotient 76 by 10....

$$76 = 7 \cdot 10 + 6$$

and finally,

$$7 = 0 \cdot 10 + 7$$

So, our series of remainders happens to be the digits of the number in base 10. Specifically $d_3 = 7$, $d_2 = 6$, $d_1 = 5$ and $d_0 = 4$.

Try doing that for 8079, but use 16 instead of 10 as the divisor.

iv) Let $\mathrm{A}_{23} = 10, \mathrm{B}_{23} = 11, \mathrm{C}_{23} = 12, \mathrm{D}_{23} = 13, \mathrm{E}_{23} = 14, \mathrm{F}_{23} = 15, \mathrm{G}_{23} = 16,$
$\mathrm{H}_{23} = 17, \mathrm{I}_{23} = 18, \mathrm{J}_{23} = 19, \mathrm{K}_{23} = 20, \mathrm{L}_{23} = 21$ and $\mathrm{M}_{23} = 22$,
then what is 185190 expressed in base-twenty-three?

v) 291480 expressed in base-twenty-three?

Answers

1. What are the following numbers expressed in base-ten?
   i) $(110101)_2 = 53$
   ii) $(\mathrm{A053D})_{16} = 656701$
   iii) $(1017)_{23} = 12197$
2. What are the following base-ten numbers expressed in an alternate base?
   i) $33 = (100001)_2$
   ii) $127 = (1111111)_2$
   iii) $8079 = (\mathrm{1F8F})_{16}$
   iv) $185190 = (\mathrm{F51H})_{23}$
   v) $291480 = (\mathrm{10M01})_{23}$

The Principle of Mathematical Induction

As we discussed way up at the top of this essay, we think about generating the set of positive integers as a process that builds them up one by one. That is, each successive integer is one more than the previous one, starting at 1, then one more taking us to 2, then 3, 4, 5, ... ad infinitum[*].

This idea of being able to step one after the other, beginning at 1 and going forever is embodied within the "Principle of Mathematical Induction" and is a basic property of the positive integers. This principle is more than just a way to generate the set of integers, it's also a way of thinking about properties of the integers.

Suppose that $P(n)$ means that the property P holds for the number n; where n is a positive integer. Then the principle of mathematical induction states that $P(n)$ is true for ALL positive integers n provided that[†]:

i) $P(1)$ is true

ii) Whenever $P(k)$ is true, $P(k+1)$ is true.

Why would these two conditions show that $P(n)$ is true for all positive integers? Note that condition ii) only asserts the truth of $P(k+1)$ under the assumption that $P(k)$ is true. However if we also know that $P(1)$ is true then condition ii) implies that $P(2)$ is true, which again implies that $P(3)$ is true, which in turn leads to the truth of $P(4)$, etc., over and over for all positive integers.

Some people picture an infinite row of dominoes. Having condition i) (called the "base case") is like being able to knock over the first domino. Then knowing condition ii) is also true is like the fact that any one domino has the ability to knock over the next. Once you've knocked over the first domino, they all fall.

Let's look at a simple example: Perhaps you've heard the story of young Carl Friedrich Gauss as a boy in the 1780s who was assigned (along with all his classmates) the tedious task of summing the first 100 integers - presumably to keep them quiet and busy while the teacher corrected some papers. Anyway, young Gauss immediately produced the answer, 5050, before most of the boys had summed the first couple of numbers. It wasn't young Gauss's extraordinary computational speed which allowed him to perform this dazzling task, but he had the deeper insight that instead of adding 1 plus 2, then adding 3, then 4, etc. he saw that if you paired 1 with 100, and 2 with 99, and 3 with 98, etc., that each of those pairs added up to 101, furthermore he knew he'd have 50 such pairs, meaning he could state the result in a heartbeat - tada - "5050"! Gauss is widely regarded as being one of the greatest mathematicians who has ever lived - the young eight-year old was just getting started.

[*] "ad infinitum" means "to infinity", or "continue forever, without limit".

[†] This wording of the definition of "The Principle of Mathematical Induction" is essentially borrowed from "Calculus" by Michael Spivak - a fabulous introductory textbook on Analysis.

Anyway, to generalize young Gauss's insight we can write the expression like this:

$$1 + 2 + 3 + \ldots + n = \frac{n(n+1)}{2}$$

So let's prove this relationship using the principle of mathematical induction.

Let $n = 1$ for the "base case", then

$$\frac{1(1+1)}{2} = \frac{2}{2} = 1$$

Which is the trivial sum* of the first positive integer 1.

Now let's assume the relationship is true for n, and prove that it must also be true for $n+1$:

$$\begin{aligned}
&(1 + 2 + 3 + \ldots + n) + (n+1) \\
&= \frac{n(n+1)}{2} + (n+1) \\
&= \frac{n(n+1)}{2} + \frac{2(n+1)}{2} \\
&= \frac{n^2 + n + 2n + 2}{2} \\
&= \frac{n^2 + 3n + 2}{2} \\
&= \frac{(n+1)(n+2)}{2} \\
&= \frac{(n+1)((n+1)+1)}{2}
\end{aligned}$$

Which proves young Gauss's expression is true for the positive integer $n + 1$ whenever it's true for n - then by the principle of mathematical induction, the expression is true for all positive integers. QED†

Intermission: Extra Exercise

If b, n are nonnegative integers and $b \neq 1$ then prove,

$$1 + b + b^2 + \cdots + b^{n-1} = \frac{b^n - 1}{b - 1}$$

Hint: use induction on n, the base case being $n = 1$.

*The word "sum" here is used in the context of the expression we are trying to prove. In this case we are summing only one item thus it's "trivial".

†"QED" - is often used at the conclusion of a proof to state that it's done - it's an acronym for the Latin phrase "quod erat demonstrandum" which means "that which was to be demonstrated". In other words we've proven what we set out to prove.

Proof for Extra Exercise

Base case: $n = 1$

$$\frac{b^1 - 1}{b - 1} = \frac{b - 1}{b - 1} = 1 = b^0 = b^{1-1}$$

Induction step: Assume the following

$$1 + b + b^2 + \cdots + b^{n-1} = \frac{b^n - 1}{b - 1}$$

Then,

$$\begin{aligned}
&(1 + b + b^2 + \cdots + b^{n-1}) + b^n \\
=&\frac{b^n - 1}{b - 1} + b^n \\
=&\frac{b^n - 1}{b - 1} + \frac{b^n(b - 1)}{b - 1} \\
=&\frac{b^n - 1 + b^{n+1} - b^n}{b - 1} \\
=&\frac{b^{n+1} + b^n - b^n - 1}{b - 1} \\
=&\frac{b^{n+1} - 1}{b - 1}
\end{aligned}$$

QED

Proof of the Basis Representation Theorem

How do we prove our theorem? There are several ways to approach it.

The mathematician George E. Andrews (in his book "Number Theory") has an interesting proof. He asks us to imagine a function that, given an integer, counts the number of base-b representations that the integer has. Then with a series of inequalities he shows that this counting function MUST produce a count of "1" for each integer, both establishing the uniqueness and the existence in one fell swoop. Cool proof - and it's short and tight, I'll bet G. H. Hardy would approve of it from an aesthetic standpoint.

We could also use the Euclidean Division Theorem to prove our theorem. We can show that by repeatedly dividing our integer n by b that we'd get a series of remainders that we could use for our base-b-digits of n. This is an interesting proof in that it also gives us a technique to construct a base-b represenation of each integer. If we were to go down this road we would try to generalize how we solved exercise number 2-iii) above, it's worth a try for fun.

However, we're going to follow a much more straightforward approach, albeit one that might not meet with G. H. Hardy's asthetic seal-of-approval as it involves looking at different "cases". This approach risks making the proof potentially tedious to follow - but we'll try to keep it light and bouncy.

The good news is that our proof will follow our intuition and experience with how we count (in base-ten) from our Sesame Street days.

Consider the odometer in your car - for every extra mile you drive the odometer keeps rolling along making a new base-ten number in its display. If your odometer stretched off to the horizon on your left, then it's pretty easy to see that there is no limit on how big a number the odometer could count to. There's no reason that our odometer has to count in base-ten either, it could be a binary odometer with only the digits one and two in each of the little dials that turn, always rolling off the miles, but showing the result in base-two. See binary odometer gif.

The idea of the odometer can be formalized by using the Principle of Mathematical Induction that we introduced earlier, which we're going to use to prove that every integer has a base-b representation.

In order to see why our imaginary odometer approach (i.e.; induction) might not lead to the "uniquness" aspect of the theorem, also imagine a magic box that also makes base-b representations for each integer - and imagine that this box makes a different representation from our odometer. It's possible! Why not - we weren't careful to show that induction is the ONLY way to generate a base-b representation of each number. In fact we just talked about another possible "magic box" that generates base-b representations of each interger - that is; we could use the Euclidean Division Theorem idea above as another way to generate base-b digits.

So - let's assume that this magic box exists and that it generates a different base-b representation for a given integer. Then we'll show that having two different ways to represent the same integer leads to a logical contradiction. So either our assumption is wrong or the axioms are. Since we're VERY confident in the axioms being correct we can only conclude that our assumption must be wrong - meaning that there is only *one* way to make a base-b represenation for each integer.

Here's a little insight into how the existence (odometer) proof works, but applied to a specific number in base-ten: All we want to show is that for some number, if you add 1 to it, that it's also possible to express it as a valid number in base-ten.

For example, adding 1 to 69412995 gives us 69412996, which is pretty trivial to show that it's valid in base-ten, only the least-significant digit was changed, and it's clearly within the range of $0 \ldots 9$.

But what about dealing with a "carry", for example if we were adding 1 to 69412999? We'd need to algebraically capture the idea of the carry. The way we do it in the proof is essentially to say that $69412999 = 69410000 + 2999 = 69410000 + (3000 - 1)$ so that when we add one to it, then it's clear that the answer is just:

$$69412999 + 1 = 69410000 + (3000 - 1) + 1 = 69410000 + 3000 + (-1 + 1) = 69413000$$

We need an algebraic mechanism like the one that rolls over the digits in the odometer.

Secondly, the magic alternate base-ten reprn machine - we will use another technique, called proof by contradiction, to prove that each such representation is unique - in other words there aren't two (or more) ways to represent the same integer in base-b. That machine is impossible, but boiled down the abstract math level.

Existence Proof of the Basis Representation Theorem

Let n and b be positive integers such that $b \geq 2$.
Also let k be a integer such that $k \geq 0$.
Furthermore let i and d_i be integers such that $0 \leq d_i \leq (b-1)$ for all $i \geq 0$.

Prove:

n can be represented as:

$$n = d_k b^k + d_{k-1} b^{k-1} + \cdots + d_2 b^2 + d_1 b^1 + d_0 b^0$$

for all n.

Proof by induction on n.

Base case:

Let $n = 1$.

Choose $k = 0$ and $d_0 = 1$. Then,

$$n = d_0 b^0 = 1 \cdot b^0 = 1 \cdot 1 = 1$$

showing that we have a valid representation for 1 in base-b since
$d_0 = 1 \leq (b-1)$, for all $b \geq 2$.

Induction Case:

Assume that n has a valid representation in base-b, that is, n can be expressed as:

$$n = d_k b^k + d_{k-1} b^{k-1} + \cdots + d_2 b^2 + d_1 b^1 + d_0 b^0$$

with all the appropriate conditions holding for d_i and k as outlined at the beginning of the proof; then we will prove that $n+1$ also has a valid representation in base-b.

We're going to break this step into two cases which cover all possibilities.

Case 1) $d_0 < (b-1)$

This case examines when the least significant digit of n is *strictly-less-than* the largest value it can take in base-b. For example, in base-two d_0 can only be zero; In base-five d_0 can be at most three; In base-ten d_0 can be at most eight, etc. This case is quite easy to deal with, so let's quickly dispense with it[*].

$$\begin{aligned} n &= d_k b^k + d_{k-1} b^{k-1} + \cdots + d_2 b^2 + d_1 b^1 + d_0 b^0 \\ &\text{if and only if,} \\ n+1 &= d_k b^k + d_{k-1} b^{k-1} + \cdots + d_2 b^2 + d_1 b^1 + d_0 b^0 + 1 \\ &= d_k b^k + d_{k-1} b^{k-1} + \cdots + d_2 b^2 + d_1 b^1 + d_0 b^0 + b^0 \\ &= d_k b^k + d_{k-1} b^{k-1} + \cdots + d_2 b^2 + d_1 b^1 + (d_0 + 1) b^0 \end{aligned}$$

we can restate our assumption that $d_0 < (b-1)$ as $d_0 \leq (b-2)$, then

$$(d_0 + 1) \leq (b-2) + 1 = (b-1)$$

showing us that the "least significant digit" of $n+1$, now being (d_0+1), is less than or equal to $(b-1)$ which means that (d_0+1) is a valid digit in base-b.

Since all the other terms $d_k, \ldots, d_2, d_1$ for $n+1$ are unchanged from their values for n then all the digits of $n+1$ are valid in base-b.

Therefore when $d_0 < (b-1)$, then $n+1$ always has a valid representation in base-b whenever n does.

Case 2) $d_0 = (b-1)$

Now we'll look at the case when the least significant digit of n is equal to the largest value it can take in base-b, that is, $d_0 = (b-1)$. (Note that "Case 1" and "Case 2" cover all the possibilities for what d_0 can be.) For example in base-two $d_0 = 1$; in base-five $d_0 = 4$; in base-ten $d_0 = 9$, etc.

Let j be the lowest power of b such that $d_j < (b-1)$, meaning we can write n as follows (…in other words all the digits to the right of d_j are equal to $(b-1)$):

$$n = d_k b^k + d_{k-1} b^{k-1} + \cdots + d_j b^j + (b-1) b^{j-1} + \cdots + (b-1) b^1 + (b-1) b^0$$

For example, if $n = 69412999$, then $j = 3$, since 10^3 is the lowest power of 10 such that its digit d_3 is less than 9 (it's 2).[†]

$$\begin{aligned} n &= d_k b^k + \cdots + d_j b^j + (b-1) b^{j-1} + (b-1) b^{j-2} + \cdots + (b-1) b^1 + (b-1) b^0 \\ &= d_k b^k + \cdots + d_j b^j + (b^j - b^{j-1}) + (b^{j-1} - b^{j-2}) + \cdots + (b^2 - b^1) + (b^1 - b^0) \\ &= d_k b^k + \cdots + (d_j b^j + b^j) + (-b^{j-1} + b^{j-1}) + \cdots + (-b^2 + b^2) + (-b^1 + b^1) - b^0 \\ &= d_k b^k + \cdots + (d_j + 1) b^j - b^0 \\ &= d_k b^k + \cdots + (d_j + 1) b^j - 1 \end{aligned}$$

[*] Recall the axiom of "Distribution" that is $a(b+c) = ab + ac$

[†] It will be helpful at this point to recall some rules of exponents, that is $a^b a^c = a^{b+c}$.

Therefore,

$$\begin{aligned} n+1 &= d_k b^k + \cdots + (d_j+1)b^j - 1 + 1 \\ &= d_k b^k + \cdots + (d_j+1)b^j \qquad (1) \end{aligned}$$

Since we picked j such that $d_j < (b-1)$, we can restate the inequality as $d_j \leq (b-2)$ therefore,

$$(d_j+1) \leq (b-2)+1 = (b-1)$$

meaning the j^{th} digit of $n+1$ is less than or equal to $b-1$ meaning that it is a valid base-b digit.

All digits $d_k, \ldots, d_{j+1}$ remain unchanged from the base-b representation of n, and all digits $d_{j-1}, \ldots, d_0$ are 0.

Therefore all the digits of the base-b representation of $n+1$ are valid in base-b.

If you've been fastidiously following the conditions on our subscript j above, then you may notice that our proof doesn't quite leave room for the case that *all* the digits are equal to $(b-1)$ because of how we defined j (we're getting picky here). For example in base-ten when $n = 99999$ there is no power of ten such that it's digit is less than 9 - remember we don't allow the most significant digit to be zero so we can't count "099999" as a valid base-ten number (if we did, we could never get our uniquness property becase $99999 = 099999 = 0000099999$).

Let's attend to this remaining detail.

Suppose $d_i = (b-1)$ for all i where $0 \leq i \leq k$, then let $d_{k+1} = 0$ and $j = k+1$.

All the arguments we just made above are essentially the same here so picking up at equation (1) above, with our new terms, we have:

$$\begin{aligned} n+1 &= (d_j+1)b^j \\ &= (d_{k+1}+1)b^{k+1} \\ &= (0+1)b^{k+1} \\ &= 1 \cdot b^{k+1} \end{aligned}$$

Meaning that $n+1$ now has one more digit than n, namely d_{k+1} and it's equal to 1, with all the rest of the digits for $n+1$ being 0 - which is a valid representation for $n+1$ in base-b for all $b \geq 2$.

Therefore when $d_0 = (b-1)$, then $n+1$ always has a valid representation in base-b whenever n does.

Taking case 1) and case 2) together proves that $n+1$ always has a valid representation in base-b whenever n does. Therefore by the principle of mathematical induction, we have proven that there is a base-b representation for all nonnegative integers.

In order to proceed with proving the uniqueness aspect of the Basis Representation Theorem, we need to make use of a well established theorem called the "Euclidean Division Theorem". It sounds onerous, but don't worry, you all learned it in the third grade but perhaps not so formally, you called it "long division". It simply states the following:

Euclidean Division Theorem[*]

For all $a, b \in \mathbb{Z}$ such that $b > 0$, there exists *unique* integers q and r such that that[†]:

$$a = qb + r \text{ such that } 0 \leq r \leq (b - 1)$$

Definition: In the above equation:

a is the *dividend*	("the number being divided")
b is the *divisor*	("the number doing the dividing")
q is the *quotient*	("the result of the division")
r is the *remainder*	("the leftover")

This is how you first learned to divide. For example if someone asks you "What is nineteen divided by three?", you'd answer "six with one remaining". Here 19 is the *dividend*, 3 is the *divisor*, 6 is the *quotient* and 1 is the *remainder*. Written in the form of the theorem:

$$19 = 6 \cdot 3 + 1$$

Often proofs make use of little mini-theorems of their own. Creating these mini-theorems is a way to simplify a step in the main proof by establishing a useful non-trivial intermediary result. It makes reading the main proof easier to follow by not having us get sidetracked with the technicalities of a step we want to make. These mini-theorems are called "Lemmas" and we're going to make one to help with proving the uniqueness part of the Basis Representation Theorem, and we're going to make use of the Euclidean Division Theorem in proving our lemma.

Lemma

Let $b, q, r \in \mathbb{Z}$ such that $b > 0$ and $0 \leq r \leq (b - 1)$, then

$$0 = qb + r$$

if and only if $q = 0$ and $r = 0$.

Proof of Lemma

Let $b, q, r \in \mathbb{Z}$ such that $b > 0$ and $0 \leq r \leq (b - 1)$.

If $q = 0$ and $r = 0$, then

$$qb + r = 0 \cdot b + 0 = 0$$

but also, by the Euclidean Division Theorem since q and r are unique for every dividend and divisor $b > 0$, then we can also conclude that if $0 = qb + r$ then $q = 0$ and $r = 0$ must be true, otherwise the quotient and remainder would not be unique. QED

[*] We introduce a new math symbol here, plus use one that was mentioned in passing much earlier in the paper. Namely the $\in$ which means "is an element of" or "is a member of" and is always followed by something that is a set. We introduced the symbol $\mathbb{Z}$ as the "set of integers". So "$a, b \in \mathbb{Z}$" means that a and b are integers.

[†] Aside: Actually the theorem is stronger than we have stated here. Specifically, it only requires that $b \neq 0$, however to keep the remainder positive, the restriction on r would have to be stated like this $0 \leq r \leq (|b| - 1)$ to deal with the possibility that b might be negative.

Uniqueness Proof of the Basis Representation Theorem

Let n and b be positive integers such that $b \geq 2$.
Also let $k \in \mathbb{Z}$ such that $k \geq 0$.
Furthermore let $i, c_i, d_i \in \mathbb{Z}$ such that $0 \leq c_i \leq (b-1)$ and $0 \leq d_i \leq (b-1)$ for all $i \geq 0$.

By the "Existence Proof of the Basis Representation Theorem" we know that n can always be expressed in this form:

$$n = d_k b^k + \cdots + d_2 b^2 + d_1 b^1 + d_0 b^0$$

Assume this expression for n is not unique and that it can also be expressed as

$$n = c_k b^k + \cdots + c_2 b^2 + c_1 b^1 + c_0 b^0$$

Let's further suppose that the index j is the lowest power such that the digits $d_j \neq c_j$ and without any loss of generality, let's assume that $d_j > c_j$.

Therefore*:

$$\begin{aligned}
& c_k b^k + \cdots + c_2 b^2 + c_1 b^1 + c_0 b^0 = d_k b^k + \cdots + d_2 b^2 + d_1 b^1 + d_0 b \\
\Leftrightarrow\quad & 0 = (d_k - c_k) b^k + \cdots + (d_j - c_j) b^j \\
\Leftrightarrow\quad & \frac{0}{b^j} = \frac{(d_k - c_k) b^k + \cdots + (d_j - c_j) b^j}{b^j}, \text{ since } b \neq 0 \\
\Leftrightarrow\quad & 0 = (d_k - c_k) b^{k-j} + \cdots + (d_{j+1} - c_{j+1}) b + (d_j - c_j) \\
\Leftrightarrow\quad & 0 = \big((d_k - c_k) b^{k-j-1} + \cdots + (d_{j+1} - c_{j+1})\big) b + (d_j - c_j)
\end{aligned}$$

Let $q = \big((d_k - c_k) b^{k-j-1} + \cdots + (d_{j+1} - c_{j+1})\big)$, then

$$0 = qb + (d_j - c_j)$$

Since $0 \leq (d_j - c_j) \leq (b-1)$ and $b > 0$ then by our lemma we know that
$q = 0$ and $d_j - c_j = 0$.

But $d_j - c_j = 0$ if and only if $d_j = c_j$ which contradicts our assumption that $d_j \neq c_j$. This implies that the initial assumption that "n is not unique" is *false*, in other words:

The base-b representation of n is unique.

Therefore since we have proven that there exists a base-b representation for ALL the nonnegative integers, *and* that this representation is unique for ALL integers, then we have proven the Basis Representation Theorem.

QED

*Please read the bidirectional arrow symbol $\Leftrightarrow$ as "if and only if" - it's like a logical "equals" sign